AFRL-RH-WP-TR-2009-0007

Predictive Validity of the Air Force Officer Qualifying Test for USAF Air Battle Manager Training Performance

Thomas R. Carretta
Warfighter Interface Division
Supervisory Control Interfaces Branch
Wright-Patterson AFB OH 45433-7511

September 2008

Interim Report for June 2008 to August 2008

Approved for public release;
Distribution is unlimited.

Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Warfighter Interface Division
Supervisory Control Interfaces Branch
Wright-Patterson AFB, OH 45433

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th ABW Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

TECHNICAL REVIEW AND APPROVAL

AFRL-RH-WP-TR-2009-0007

THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION.

FOR THE DIRECTOR

//**signed//**	**//signed//**
Thomas R. Carretta	Daniel G. Goddard
Research Engineering Psychologist	Chief, Warfighter Interface Division
Supervisory Control Interfaces Branch	Human Effectiveness Directorate
	711th Human Performance Wing

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*
02 Sep 2008

2. REPORT TYPE
Interim

3. DATES COVERED *(From - To)*
June 2008 – August 2008

4. TITLE AND SUBTITLE
Predictive Validity of the Air Force Officer Qualifying Test for USAF Air Battle Manager Training Performance

5a. CONTRACT NUMBER
IN-HOUSE

5b. GRANT NUMBER
NA

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S
Thomas R. Carretta

5d. PROJECT NUMBER
2313

5e. TASK NUMBER
HC

5f. WORK UNIT NUMBER
2313HC58

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Material Command
Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Interface Division
Wright-Patterson AFB OH 45433

8. PERFORMING ORGANIZATION REPORT NUMBER
NA

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Material Command
Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Warfighter Interface Division
Supervisory Control Interfaces Branch
Wright-Patterson AFB OH 45433

10. SPONSOR/MONITOR'S ACRONYM(S)
711 HPW/RHCI

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-RH-WP-TR-2009-0007

12. DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release; Distribution is unlimited.

13. SUPPLEMENTARY NOTES
88 ABW PA Cleared 09/18/2008, WPAFB-08-0156.

14. ABSTRACT
The Air Force Officer Qualifying Test (AFOQT) was validated for the prediction of academic performance during Undergraduate Air Battle Manager (UABM) training. Participants were 680 men and women selected for UABM training on the basis of academic performance, aptitude test scores (i.e., AFOQT), commander's ratings, and medical and physical fitness. The criterion was the mean on several written tests taken during the 9 month UABM training course. Validities were examined for the AFOQT subtests, operational composites, and content-based composites in both observed form and after correction for range restriction. All of the subtests and composites were significantly related to academic performance during training. Results of the subtest-level and content-based composite analyses provided insights as to the development of a UABM qualification composite. Prior to implementation, additional studies are required to examine predictive validity against other performance criteria (e.g., pass/fail training, hands-on work samples). Other studies are needed to determine the effects of alternate composites on sex and racial/ethnic subgroup performance including qualification rates (i.e., adverse impact) and predictive validity (i.e., predictive bias).

15. SUBJECT TERMS
Undergraduate Air Battle Manager training, Air Force Officer Qualifying Test, predictive validation

16. SECURITY CLASSIFICATION OF: Unclassified			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Thomas R. Carretta
U	U	U	UNL	25	19b. TELEPONE NUMBER (*Include area code*)

THIS PAGE INTENTIONALLY LEFT BLANK.

CONTENTS

TABLES

PREFACE

This report describes activities performed in support of USAF personnel selection and classification (AF/A1PF), Work Unit 2313HC58. The author thanks Mr Ken Schwartz (AETC/DPSF) and the AFPC Human Resources Data Bank (HRRD) for support in the development of the database used in this study.

THIS PAGE IS INTENTIONALLY LEFT BLANK.

INTRODUCTION

The Air Force Office Qualifying Test (AFOQT; Carretta & Ree, 1996) has been used by the Officer Training School and Reserve Officer Training Corps for officer commissioning and aircrew training qualification since 1957. The AFOQT has been validated against officer training performance (Roberts & Skinner, 1996), several pilot and navigator (combat system operator) training performance criteria including passing/failing training, training grades, class rank (Carretta & Ree, 2003; Olea & Ree, 1994), and several non-aviation officer jobs (Arth, 1986; Arth & Skinner, 1986; Finegold & Rogers, 1985). Despite its long usage by the US Air Force (USAF), no studies were found that examined its predictive validity for air battle manager training. The purpose of this study was to examine its predictiveness against Undergraduate Air Battle Manager training performance.

What are Air Battle Managers?

The USAF air battle manager (ABM) specialty has been a rated career field since 1999. ABMs are primarily concerned with command and control, ensuring the execution of daily air missions. ABMs control the air battlespace, aid the flight, and keep track of all assets in the area of operations to ensure deconfliction and safety. Their specific duties vary, depending on the overall military operation. For air-to-air combat missions, ABMs use radar to provide the pilot and other aircrew the information they need to increase situational awareness and to find, identify, and prosecute targets. ABMs can provide early warning regarding inbound enemy aircraft and direct friendly assets to intercept them (Miller, 1997).

Undergraduate Air Battle Manager (UABM) training for the active duty Air Force, Air National Guard, and Air Force Reserve is conducted at Tyndall AFB, FL by the 325th Air Control Squadron. The focus of the 9 month course is to orient new officers to battle management concepts and the duties they will perform at their first unit as a battle management apprentice. Course content includes air surveillance concepts, basic controlling techniques, weapon systems capabilities and tactics, and an introduction to integrated (weapons and surveillance) battle management. Upon completion of UABM

training, graduates enter unit-level upgrade (apprentice) training where the focus is on training controller skills and expanding the knowledge base. Battle management apprentice training involves from 18-36 months of orientation-level weapon controller training. During this period, ABMs transition to their first battle management position either as a senior director or an air surveillance officer. Additional intermediate and advanced training opportunities are available for experienced controllers either on-the-job or through specialized training courses (e.g., Counter Air Tactics Awareness Training Course at Tyndall AFB, FL; USAF Weapons School at Nellis AFB, NV).

UABM training applicants must first qualify for an officer commissioning program through the United States Air Force Academy (USAFA), Officer Training School (OTS), or the Reserve Officer Training Corps (ROTC). UABM training candidates are evaluated for training suitability by selection boards within each commissioning program. As with other officer aircrew specialties (pilot, combat system operator), ABM training selection factors include academic performance, aptitude test scores, commander’s ratings, and medical and physical fitness US Air Force, 2006).

Purpose

The purpose of this study was to examine the predictiveness of AFOQT subtests and composites against UABM training. The operational AFOQT composites were used as a baseline. Alternate composites also were evaluated.

METHODS

Participants

The sample consisted of 680 students who had tested on the AFOQT between 1999 and 2007 and subsequently attended Undergraduate Air Battle Manager (UABM) training. The sample was mostly male (82.1%) and white (79.1%). In addition to qualification on the AFOQT, officer commissioning and aircrew training applicants met

various academic (e.g., college degree), fitness (e.g., physical fitness test), medical (e.g., physical exam), moral (e.g., legal issues), and physical (e.g., weight) standards.

Measures

Participants tested on the AFOQT between 1999 and 2007. As a consequence, most participants (94.4%) tested on AFOQT Form Q, which consisted of 16 cognitive subtests. When AFOQT Form S was implemented in July 2005, five of the subtests from previous forms (O, P, and Q) had been removed. AFOQT Form S consists of 11 cognitive subtests that are combined into five composites (see Table 1). For the purpose of this study, AFOQT raw score composites were computed on the basis of the Form S content and composite specifications. Personnel decisions including qualification for officer commissioning programs and aircrew training are made, in part, on the basis of the composites. Brief descriptions of the AFOQT subtests grouped by content are presented below.

Verbal subtests. Verbal Analogies (VA) provide a measure of the ability to reason and determine relationships between words. Word Knowledge (WK) assesses verbal comprehension involving the ability to understand written language through the use of synonyms.

Quantitative subtests. Arithmetic Reasoning (AR) measures the ability to understand arithmetic relations expressed as word problems. Math Knowledge (MK) provides a measure of the ability to use mathematical terms, formulas, and relations.

Spatial subtests. Block Counting (BC) measures spatial ability through the analysis of three-dimensional representations of a set of blocks. Rotated Blocks (RB) assess the ability to visualize and mentally manipulate objects. Hidden Figures (HF) measure the ability to see a simple figure embedded in a complex drawing.

Table 1. Composition of AFOQT Form S Aptitude Composites

	Composite				
Subtest	Verbal (V)	Quantitative (Q)	Academic Aptitude (AA)	Pilot (P)	Navigator/ Technical (N/T)
Verbal Analogies (VA)	X		X		X
Arithmetic Reasoning (AR)		X	X	X	X
Word Knowledge (WK)	X		X		
Math Knowledge (MK)		X	X	X	X
Instrument Comprehension (IC)				X	
Block Counting (BC)					X
Table Reading (TR)				X	X
Aviation Information (AI)				X	
Rotated Blocks (RB)					
General Science (GS)					X
Hidden Figures (HF)					

Note. Although RB and HF were retained in AFOQT Form S, they do not contribute to any of the operational composites.

Aircrew subtests. Instrument Comprehension (IC) assesses the ability to determine the attitude of an aircraft from illustrations of flight instruments. Aviation Information (AI) measures knowledge of general aviation terms, concepts, and principles. General Science (GS) provides a measure of knowledge and understanding of scientific, terms, concepts, instruments, and principles.

Perceptual speed subtests. Table Reading (TR) assesses the ability to quickly and accurately extract information from tables.

Table 2 summarizes the minimum qualifying AFOQT scores for USAF officer commissioning and aircrew training programs as described in Air Force Instruction 36-2013 (United States Air Force, 2006). Though the same composites are used to qualify applicants for pilot, navigator (control system operator), and air battle manager training, the minimum qualifying scores vary among the programs.

Table 2. AFOQT Category Minimum Qualification Scores

Program	V	Q	P	N/T	Other
Officer Commissioning	15	10	--	--	--
Flying Training (Pilot or Helicopter)	15	10	25	10	(P + N) ≥ 50
Flying Training (Navigator)	15	10	10	25	(P + N) ≥ 50
Flying Training (Air Battle Manager)	15	10	10	10	(P + N) ≥ 50

Note. Verbal = V, Quantitative = Q, Pilot = P, and Navigator/Technical = N/T.

Criterion

Undergraduate Air Battle Manager final score (UABM FS) represents mean performance on several written tests taken during the 9 month training course at Tyndall

AFB, FL. Students must achieve a score of 85 or higher on each written test to receive a passing score for the course. Students are allowed one test and one retest. UABM FS was available only for course graduates. UABM training attrition is normally about 10%.

Procedures

Data were collected from official personnel records. No sampling was used in this study. All air battle manager trainees with complete AFOQT scores and UABM training outcome data were included. Test score data were collected prior to entrance into USAF officer commissioning programs and UABM training. As previously noted, the study included students who tested on two different forms of the AFOQT (Forms Q and S). Although the same composite scores are computed for each form, their subtest composition differs. AFOQT raw score composites were calculated by summing the subtests on the basis of the Form S content and composite specifications. Training performance data were collected during the 9 month UABM training program.

As this is a range restricted sample due to prior selection into US Air Force officer commissioning and UABM training, the multivariate correction for range restriction (Lawley, 1943; Ree, Carretta, Earles, & Albert, 1994) was applied. The unrestricted estimates of means, variances, and correlations for the AFOQT subtests came from USAF officer applicants. The multivariate procedure estimates the means, variances, and correlations of the predictors and criterion as they would be found in the unrestricted population.

Analyses

Analyses were performed at both the AFOQT subtest and composite score level. To begin, descriptive statistics and correlations between the AFOQT scores and UABM training final score (UABM FS) were examined. Next, the UABM final score was regressed on the AFOQT scores. Additional analyses were performed to evaluate alternate AFOQT composites. All analyses used a .05 Type I error rate.

RESULTS

Descriptive Statistics

Table 3 shows the means and standard deviations for the AFOQT subtests and for the UABM final score for the observed and corrected data. On average, the restriction in

Table 3. Means and Standard Deviations for AFOQT Subtests and UABM Training Final Score

	Observed Data		Corrected Data	
Score	Mean	SD	Mean	SD
VA	17.58	3.55	16.55	4.27
AR	15.39	4.87	15.69	5.45
WK	16.49	4.68	15.87	5.46
MK	16.22	4.80	15.93	5.48
IC	13.11	4.85	12.76	5.34
BC	13.68	3.62	12.24	4.12
TR	29.03	6.42	27.02	6.60
AI	9.26	4.22	10.52	4.53
RB	8.99	2.98	9.70	3.40
GS	10.48	3.44	13.06	3.82
HF	10.26	3.17	10.58	3.50

UABM FS	94.43	2.47	94.57	2.61

N = 680

range was such that the variances in the sample were about 80% of the population variance values. This reduction in variance biases the correlation estimates. The magnitude of the bias is a function of the degree to which the variances were curtailed.

Correlations

Subtest analyses. The observed correlations of the AFOQT subtests and the criterion are presented above the diagonal in Table 4. All correlations, except that between MK and AI ($r = .023$, ns) were statistically significant at the $p < .01$ Type I error rate. All of the AFOQT subtest scores were significantly correlated with the training criterion. The mean observed correlation between the 11 subtests and the UABM training final score was .205 and ranged from .117 (HF) to .271 (WK).

The correlations corrected for range restriction are presented below the diagonal. After correction, the mean correlation between the 11 AFOQT subtests and the UABM training final score was .295 and ranged from .201 (HF) to .356 (AR).

Composite score analyses. Table 5 displays the correlations among the AFOQT composites and UABM FS training criterion. The observed correlations are presented above the diagonal. All observed correlations were statistically significant at the .01 Type I error rate. The mean observed correlation between the 5 AFOQT composites and the UABM final score was .311 and ranged from .274 (Quantitative) to .338 (Pilot).

The correlations after correction for range restriction are presented below the diagonal. After correction, the mean correlation between the 5 AFOQT composites and the training criterion was .411 and ranged from .376 (Quantitative) to .447 (Academic Aptitude).

Table 4. Correlation Matrix of the AFOQT Subtests and the UABM Final Score Criterion

Score	VA	AR	WK	MK	IC	BC	TR	AI	RB	GS	HF	UABM FS
VA	1.000	0.430	0.646	0.292	0.217	0.239	0.146	0.165	0.287	0.444	0.269	0.246
AR	0.513	1.000	0.352	0.611	0.240	0.342	0.353	0.128	0.335	0.446	0.310	0.268
WK	0.691	0.445	1.000	0.237	0.181	0.209	0.103	0.214	0.218	0.490	0.199	0.271
MK	0.422	0.706	0.358	1.000	0.168	0.285	0.279	0.023	0.245	0.405	0.236	0.224
IC	0.376	0.455	0.310	0.399	1.000	0.351	0.243	0.469	0.410	0.286	0.375	0.176
BC	0.351	0.484	0.271	0.410	0.508	1.000	0.470	0.129	0.384	0.256	0.372	0.138
TR	0.249	0.388	0.182	0.310	0.332	0.486	1.000	0.105	0.212	0.151	0.265	0.203
AI	0.371	0.349	0.363	0.281	0.611	0.357	0.250	1.000	0.246	0.328	0.184	0.224
RB	0.353	0.478	0.285	0.423	0.563	0.507	0.307	0.418	1.000	0.344	0.443	0.176
GS	0.563	0.530	0.549	0.566	0.470	0.361	0.211	0.485	0.443	1.000	0.295	0.216
HF	0.348	0.421	0.270	0.395	0.485	0.492	0.335	0.333	0.543	0.376	1.000	0.117
UABM FS	0.339	0.356	0.354	0.320	0.286	0.228	0.258	0.327	0.266	0.309	0.201	1.000

Note. Correlations above the diagonal are observed. Those below the diagonal were corrected for range restriction (Lawley, 1943).

N = 680

Table 5. Correlation Matrix of the AFOQT Composites and the UABM Final Score Criterion

Score	V	Q	AA	P	N/T	UABM FS
V	1.000	0.397	0.808	0.385	0.558	0.286
Q	0.506	1.000	0.861	0.765	0.844	0.274
AA	0.850	0.884	1.000	0.704	0.850	0.334
P	0.523	0.836	0.793	1.000	0.880	0.338
N/T	0.653	0.878	0.889	0.925	1.000	0.321
UABM FS	0.378	0.366	0.428	0.424	0.415	1.000

Notes. AFOQT Verbal = V, AFOQT Quantitative = Q, AFOQT Academic Aptitude = AA, AFOQT Pilot = P, AFOQT Navigator/Technical = N/T, and Undergraduate Air Battle Manager Final Score = UABM FS. Correlations above the diagonal are observed. Those below the diagonal were corrected for range restriction (Lawley, 1943).

N = 680

Test content composite score analyses. In addition to examining the predictive validity of the AFOQT subtests and operational composites, we thought it might be informative to examine validity as a function of test content. Factor analyses of the AFOQT (Carretta & Ree, 1996; Drasgow, Nye, Carretta, & Ree, in press; Skinner & Ree, 1987) have identified 5 content areas: verbal (VA, WK), math (AR, MK), spatial (BC, RB, HF), aircrew (IC, AI, GS), and perceptual speed (TR). Unit-weighted composites[1] were computed for these content areas and correlated with UABM final score. The observed correlations of these content-related composites with the criterion were: Verbal (.286), Math (.274), Spatial (.184), Aircrew (.265), and Perceptual Speed (.203). After

[1] It should be noted that the Verbal and Math composites are identical to the AFOQT Verbal and Quantitative composites and that the Perceptual Speed composite is composed only of the TR subtest.

correction for range restriction the correlations were: Verbal (.378), Math (.366), Spatial (.281), Aircrew (.369), and Perceptual Speed (.258).

Regression Analyses

Several regression models were evaluated to determine the predictiveness of the AFOQT subtests and composites versus UABM training final score.

Subtest analyses. The subtest level regression analyses are summarized in Table 6. Model 1 included all 11 AFOQT subtests. Although the model was statistically significant ($R = .394$, $p < .01$), examination of the beta weights indicated that all 3 of the spatial ability subtests (BC, RB, HF) had negative, non-significant weights, so these scores were removed and another regression was performed. Model 2 had 8 scores, was statistically significant ($R = .390$, $p < .01$), and did not differ from the starting model with 11 scores ($F(3, 668) = 0.91$, ns). Examination of the beta weights for Model 2 indicated that 2 of the aviation-related subtests (IC, GS) had non-significant weights so they were removed and another regression done. Model 3 had 6 scores including both verbal (VA, WK) and both math (AR, MK) subtests, AI, and TR. It also was statistically significant ($R = .389$, $p < .01$) and did not differ from Model 2 ($F(2, 671) = 0.24$, ns). Examination of the beta weights revealed that they were non-significant for one of the verbal (VA) and math (AR) subtests. It is likely that the non-significant weights were due to the strong correlations between VA and WK ($r = .646$, .691 after correction) and between AR and MK ($r = .611$, .706 after correction). See Table 4. Subsequent models examined this hypothesis through the use of composite scores instead of subtests. The fourth and final subtest model had 4 scores (WK, MK, AI, TR) after the removal of VA and AR. As with the other models, it was statistically significant ($R = .381$, $p < .01$) and did not significantly differ from Model 3 in predictiveness ($F(2, 673) = 2.32$, ns). After correction for range restriction, the correlation of Model 4 with UABM final score increased to .464.

Table 6. Summary of Subtest-Level Regression Analyses

Model	N Scores	R	R^2	R^2 Change	F Change	df1	df2
1	11	.394**	.155				
2	8	.390**	.152	-.003	0.91	3	668
3	6	.389**	.151	-.001	0.24	2	671
4	4	.381**	.146	-.006	2.32	2	673

N = 680. **p <.01

Composite score analyses. The AFOQT operational composite regression analyses are summarized in Table 7. Model 1 included the AFOQT Verbal (V), Quantitative (Q), Pilot (P), and Navigator/Technical (N/T) composites. Academic Aptitude (AA) was not included as it is redundant with V and Q. The rationale for this model is that all of these composites contribute to UABM training qualification; V and Q for officer commissioning and P and N/T for UABM training qualification. Model 1 was statistically significant (R = 381, p < .01), but included a score with a negative beta weight (N/T). Model 2 removed the N/T composite (R = .378, p < .01) with little loss in predictiveness (F (1, 675) = 1.67, ns). Examination of the Model 2 beta weights indicated that the Q composite was non-significant. This was not surprising as the 2 math subtests (AR, MK) that comprise the Q composite also are included in the P composite. Model 3 included only the V and P composites. It was statistically significant (R = .378, p < .01) and did not differ from Model 2 (F (1, 676) = 0.01, ns). The correlation of Model 3 (V and P composites) with UABM final score increased to .462 after correction for range restriction.

Table 7. Summary of AFOQT Composite Regression Analyses

Model	N Scores	R	R^2	R^2 Change	F Change	df1	df2
1	4	.381**	.145				
2	3	.378**	.143	-.002	1.67	1	675
3	2	.378**	.143	.000	0.01	1	676

N = 680. **p <.01

Test content composite score analyses. The content-based composite regression analyses are summarized in Table 8. Model 1 included the 5 content-based composites: verbal (VA, WK), math (AR, MK), spatial (BC, RB HF), aviation (IC, AI, GS), and perceptual speed (TR). The advantage of this approach over using the operational composites is that the content-based composites are not redundant (no overlapping subtest scores). Model 1, with all 5 content-based composites, was statistically significant (R = .377, p < .01). Examination of the beta weights indicated that the spatial content composite was not significant, so it was removed and the analyses repeated. The resulting model, Model 2, was statistically significant (R = .375, p < .01), all of the regression weights were positive and significant, and the model did not differ in predictiveness from the previous model (F (1, 674) = 1.54, ns). We next computed a unit-weighted composite that combined the 8 subtests that contributed to the verbal, math, aviation, and perceptual speed content-based composites (VA, WK, AR, MK, IC, AI, GS, TR). The resulting unit-weighted composite had a correlation of .373 with the UABM final score, showing little loss of predictiveness from the regression-weighted content-based composites (.373 vs. .375). After correction for range restriction the unit-weighted composite correlated .457 with the UABM final score.

Table 8. Summary of Content-Based Composite Regression Analyses

Model	N Scores	R	R^2	R^2 Change	F Change	df1	df2
1	5	.377**	.142				
2	3	.375**	.140	-.002	1.54	1	674
3	1	.373**	.139				

Note. Model 3 is a unit-weighted composite of the subtest scores that contribute to the Model 2 content-level composites. R^2 Change could not be computed for Model 3.

N = 680. **p < .01

DISCUSSION

Correlational analyses indicated that the AFOQT subtests were related significantly with academic performance during UABM training. These results are consistent with prior validations of the AFOQT versus academic performance criteria in pilot (Carretta & Ree, 1995; Olea & Ree, 1994; Ree, Carretta, & Teachout, 1995)) and navigator (Olea & Ree, 1994) training.

Subsequent analyses took three different approaches to examine the predictiveness of the subtests when used in combination. These included examination of the subtests, operational composites, and content-based composites.

The analyses based on the operational composites should be treated as a baseline, since they are used to make personnel selection and classification decisions. UABM applicants must qualify for an officer commissioning program based on the Verbal and Quantitative composites, and then qualify for UABM training based on the Pilot and Navigator/Technical composites. The regression analyses using the operational composites revealed that only the Verbal and Pilot composites contributed to the prediction of UABM training performance. This was likely due to overlap in content

among the operational composites (see Table 1). The 2 math subtests (AR, MK) contribute to the Q, P, and N/T composites. VA contributes to V and N/T and TR contributes to P and N/T.

While the operational composite analyses were informative regarding the predictive utility of the AFOQT versus UABM training performance as it currently is used, the subtest-level and content-level analyses provided insights as to the development of UAMB training qualification composite. In both approaches, the spatial content subtests (BC, RB, and HF) did not provide incremental validity when used with the other subtests. Further, 2 of the aviation content subtests (IC, GS) did not provide incremental validity in the subtest-level regression analyses. The only subtests that contributed in all 3 approaches were WK (subtests, Verbal composite, verbal content composite), MK (subtest, Pilot composite, math content composite), AI (subtests, Pilot composite, aviation content composite), and TR (subtests, Pilot composite, perceptual speed content composite). The multiple R for a regression-weighted composite of these 4 subtests (WK, MK, AI, TR) was .381 for the observed data and .464 for the corrected data. A unit-weighted composite of these 4 subtests had only slightly lower validity than the regression-weighted subtests (r of .371 and .456 for the observed and corrected data).

Although VA and AR did not provide *incremental* validity in the subtest-based regressions, they were among the 4 highest subtest validities after correction for range restriction. A regression-weighted composite consisting of VA, AR, WK, MK, AI, and TR had observed and corrected validities of .389 and .470 respectively. The validity of a unit-weighted composite of these 6 subtests was only about .01 lower than the regression-weighted values (observed and corrected validities of .379 and .463).

In sum, although the operational AFOQT composites demonstrated acceptable validity versus training performance, results of the correlational and regression analyses indicated that UABM training selection could improve if an alternate composite were used. However, additional studies are required prior to implementation of a new composite. To begin, it would be informative to conduct predictive validation studies with alternate performance criteria such as passing/failing training and on a measure of hands-on job performance based on work samples similar to check ride scores for pilot

training. Other studies are needed to determine the effects of alternate composites on setting minimum qualifying scores and on sex and racial/ethnic subgroup performance including qualification rates (i.e., adverse impact) and predictive validity (i.e., predictive bias).

REFERENCES

Arth, T. O. (1986). *Validation of the AFOQT for non-rated officers,* AFHRL-TP-85-50. Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Arth, T. O., & Skinner, J. (1986). *Aptitude selection for Air Force officer non-aircrew jobs.* Paper presented at the annual meeting of the Military Testing Association, Mystic, CT.

Carretta, T. R., & Ree, M. J. (1995). Air Force Officer Qualifying Test Validity for predicting pilot training performance. *Journal of Business and Psychology, 9,* 379-388.

Carretta, T. R., & Ree, M. J. (1996). Factor structure of the Air Force Officer Qualifying Test: Analysis and comparison. *Military Psychology, 8,* 29-42.

Carretta, T. R., & Ree, M. J. (2003). Pilot selection methods. In B. H. Kantowitz (Series Ed.) & P. S. Tsang & M. A. Vidulich (Vol. Eds.). *Human factors in transportation: Principles and practices of aviation psychology* (pp. 357-396). Mahwah, NJ: Erlbaum.

Drasgow, F., Nye, C. D., Carretta, T. R., & Ree, M. J. (in press). *Factor structure of the Air Force Officer Qualifying Test form S: Analysis and comparison with previous forms,* AFRL-HE-WP-TR-2008-xxxx. Wright-Patterson AFB, OH: Air Force Research Laboratory, Warfighter Interfaces Division.

Finegold, L., & Rogers, D. (1985). *Relationship between Air Force Officer Qualifying Test scores and success in air weapons controller training,* AFHRL-TR-85-13.Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Lawley, D. N. (1943). A note on Karl Pearson’s selection formulae. *Proceedings of the*

Royal Society of Edinburgh, Section A, 62, Part 1, 28-30.

Miller, C. B. (1997). *USAF TACS battle management: Preparing for high tempo future operations.* Master's Thesis submitted to Marine Corps University, Quantico, VA.

Olea, M. M., & Ree, M. J. (1994). Predicting pilot and navigator criteria: Not much more than *g*. *Journal of Applied Psychology, 79,* 845-849.

Ree, M. J., Carretta, T. R., Earles, J. A., & Albert, W. (1994). Sign changes when correcting for range restriction: A note on Pearson's and Lawley's selection formulas. *Journal of Applied Psychology, 79,* 298-301.

Ree, M. J., Carretta, T. R., & Teachout, M. S. (1995). Role of ability and prior job knowledge in complex training performance. *Journal of Applied Psychology, 80*, 721-730.

Roberts, H. E., & Skinner, J. (1996). Gender and racial equity of the Air Force Officer Qualifying Test in officer training school selection decisions. *Military Psychology, 8,* 95-113.

Skinner, J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and factor analysis of Form O,* AFHRL-TR-86-68. Brooks Air Force Base, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

United States Air Force (2006). *Personnel: Officer Training School and Enlisted Commissioning Programs,* Air Force Instruction 36-2013. Washington, DC: Secretary of the Air Force.